Journal of Vegetation Science 14: 899-906, 2003
© IAVS; Opulus Press Uppsala.

Effects of sampling teams and estimation methods on the assessment of plant cover

Kercher, Suzanne M.[1]; Frieswyk, Christin B.[1,2] & Zedler, Joy B.[1*]

[1]Department of Botany, University of Wisconsin-Madison, 430 Lincoln Drive, Madison, WI, USA 53706;
E-mail: skercher@wisc.edu; [2]cbfrieswyk@wisc.edu
[*]Corresponding author; Fax+16082627509; E-mail jbzedler@wisc.edu

Abstract. We evaluated variability in cover estimation data obtained by (1) two sampling teams who double sampled plots and (2) one team that used two methods (line intercepts and visual estimation of cover classes) to characterize vegetation of herbaceous wetlands. Species richness and cover estimates were similar among teams and among methods, but one sampling team scored cover higher than the other. The line intercept technique yielded higher cover estimates but lower species richness estimates than the cover class method. Cluster analyses of plots revealed that 36% and 11% of plots sampled consecutively by two teams or using two methods, respectively, were similar enough in species composition and abundance to be paired together in the resulting clustering tree. Simplifying cover estimate data to presence/absence increased the similarity among both teams and methods at the plot scale. Teams were very similar in their overall characterization of sites when cover estimation data were used, as assessed by cluster analysis, but methods agreed best on their overall characterization of sites when only presence/absence data were considered. Differences in abundance estimates as well as pseudoturnover contribute to variability. For double sampled plots, pseudoturnover was 19.1%, but 57.7% of pseudoturnover cases involved taxa with ≤ 0.5% cover while only 3.4% involved taxa with > 8% cover. We suggest that vegetation scientists incorporate quality control, calibrate observers and publish their results.

Keywords: Cover class; Line intercept; Observer bias; Observer error; Pseudoturnover; Quality control; Sampling error; Wisconsin.

Introduction

Scientists are increasingly held to high standards for the quality of their data by both reviewers and funding agencies (Anon. 1998). The goal of quality control is to quantify and understand variation and ensure that data meet defined standards of quality, not to eliminate error or variation in the data (Anon. 1996). Despite increasing requests for scientists to quantify and control error in their data, few examples of quality control have appeared in vegetation science literature. Quality control in vegetation sampling requires that sources of variation and the contribution of those sources to the overall variation of the vegetation sample are identified. Of the two general components of variation in data, the accuracy of the sample is of obvious importance but is very difficult to assess, particularly in large-scale field surveys under typical constraints of labour and time shortages. The precision of the sample involves the degree of similarity among measured values without regard to accuracy (Gotfryd & Hansell 1985).

Vegetation sampling precision may be influenced by different observers and type of sampling method used (e.g. cover class vs line intercept). The effect of different observers is generally believed to be an important contributor to variability in the data set (Smith 1944; Lamacraft 1978; Sykes et al. 1983; Gotfryd & Hansell 1985; Nilsson & Nilsson 1985; Kirby et al. 1986; West & Hatton 1990; Lepš & Hadincová 1992; Rich & Woodruff 1992; Rich & Smith 1996; Klimeš et al. 2001), though Oredsson (2000) reported negligible observer error compared to other errors. Likewise, there may be variability attributable to the sampling method(s) chosen by researchers (e.g. Kirby et al. 1986; Rich & Woodruff 1992). Some scientists find reasons to doubt the efficacy of particular methods (e.g. Smith 1944; Guo & Rundel 1997; Oredsson 2000). For example, in a newsletter of the International Association of Vegetation Scientists, Wilson (unpubl.) equated estimating cover by eye to 'guessing cover' and expressed doubt that even a line intercept method could assess cover accurately in a stand of mixed herbaceous species. He concluded that presence/absence data are the only acceptable alternative. Others argue that even subjective measures, such as cover, are valuable because they add extra information (e.g. Sykes et al. 1983).

The need for quality control of vegetation surveys has been recognized since at least the 1940s (Hope-Simpson 1940; Smith 1944). Considering the temporal limitations, shortages of labour and differences in observer expertise and experience in most large-scale field surveys, the challenge for plant ecologists is to find vegetation sampling methods that are robust (repeatable) across observers and yet efficient in the effort required to characterize the vegetation adequately. We present results from two field research projects in Wisconsin, USA. Each was designed to characterize the species composition and cover within temperate wetland communities dominated by herbaceous grasses, sedges and forbs. We asked:

1. How do sampling teams affect the reporting of species composition and cover? Are biases detectable and, if so, are they attributable to the growth form of the plants?
2. How well does a cover class estimation method match a line intercept sampling method? Are broad cover classes more repeatable than fine classes?

Methods

Two data sets from two separate studies were used. We compared two sampling teams in the first project (southern Wisconsin wet meadows) and two sampling methods in the second project (Lake Michigan wetlands).

Data set 1: Sampling team variability

Twelve wet meadows in Dane County, Wisconsin were sampled by two teams from 10.07-20.07.2000. A sampling team consisted of two trained field botanists. All four botanists were graduate students in Botany or Environmental Studies, and all four had completed an intensive three week summer field course in wetland plant identification in southern Wisconsin.

Ten 1 m × 1 m plots were sampled consecutively ('double sampled') by the two teams (hereafter called Team A and Team B) on each site, for a total of 120 double sampled plots across the 12 sites. Every 9 m along a 90 m transect line, a team would walk a random distance of 1 - 25 m away from and perpendicular to the transect line and place a three sided square PVC frame on the ground. After being sampled by the first team, the corners of the plots were marked so that the second team could locate them.

Within a plot, species were assigned to one of nine cover classes in the $\log_2$ system (Gauch 1982):
1 = ≤ 0.5%; 2 = 0.5-1%; 3 = 1-2%; 4 = 2-4%; 5 = 4-8%;
6 = 8-16%; 7 = 16-32%; 8 = 32-64%; 9 = 64-100%.
Small PVC frames representing the different cover classes were available for use as needed to 'calibrate' the sampling teams. Unknown taxa were collected, pressed, dried and identified by a professional taxonomist (T. Cochrane) at the University of Wisconsin-Madison Herbarium.

Data set 2: Comparison of two vegetation sampling methods

Two methods for estimating plant abundance were compared as part of the Great Lakes Environmental Indicators project. This data set was collected from eight coastal wetlands on the coast of Green Bay and nearby Lake Michigan, Wisconsin, USA. A total of 270 plots were sampled across the eight wetlands; the number of plots sampled on each site was proportional to the area of the wetland.

In each wetland 1 m × 1 m plots were placed in a stratified random design along randomly placed transects that ran nominally perpendicular to the water gradient. Each plot was randomly located within every 20-m interval of the transect, with the exception of a non-random plot that was placed at the wet end of the transect. Plots were situated directly adjacent to the transect line. Within plots, the abundance of each plant species was estimated using a cover class method and a line intercept method. The cover class system was a modified Braun-Blanquet method using six cover categories:
1 = <1%; 2 = 1-5%; 3 = 5-25%; 4 = 25-50%; 5 = 50-75%;
6 = 75-100%.
For the line intercept method, two 1-m sticks were placed within the plot 25 cm and 75 cm from one side and perpendicular to the transect. Each stick was divided into 10 cm intervals and the number of intervals each plant species intercepted was recorded.

Prior to analysis of data set 2, cover class and line intercept data were transformed into percentages. This was done by taking the midpoint of each cover class and the percent of line intervals crossed by a species, respectively.

Data analyses

For data set 1, we compared the species richness values and sums of cover classes reported by the two teams for the double sampled plots using paired *t*-tests and correlation coefficients (*r*). To compare the teams' reporting of the species composition and abundances within plots, the ten double sampled plots from each of the 12 sites were subjected to a divisive hierarchical cluster analysis using non-standardized Euclidean distances in S-plus 2000 (Anon. 2000). Prior to all cluster analyses, cover class data were transformed into percentages by taking the midpoint of each cover class. Plots were first clustered using the $\log_2$ system. The analysis was repeated three times, once with cover classes reduced to six categories:

1 = < 4%; 2 = 4-8%; 3 = 8-16%; 4 = 16-32%; 5 = 32-64%; 6 = 64-100%;

then with three cover categories:

1 = < 8%; 2 = 8-32%; 3 = 32-100%;

and finally with presence/absence categories only.

For data set 2, paired *t*-tests and correlation coefficients were calculated to compare the species richness and species abundance estimates for (1) one vs two lines, (2) one line vs cover class and (3) two lines vs cover class. Divisive hierarchical cluster analyses were also used to compare vegetation sampling methods, again considering each of the eight sites separately. In both cluster analyses, the plots that were sampled twice by different teams or by different methods were considered to be 'paired correctly' in the clustering tree if they occupied the same ultimate branch.

Jaccard's and Horn's indices (Brower et al. 1990) were used to assess the similarity of results reported by two sampling teams (using data set 1) and sampling methods (using data set 2) at three different analytical scales. To assess overall similarity, we calculated the mean percent cover of all species sampled by each team across all 12 sites. To compare methods we calculated mean percent cover and frequency values for each species sampled across all eight sites. On the scale of sites, we calculated the similarity among teams or methods by tabulating the mean percent cover values for all species reported from each team or method on a site, and then allowed the differences within each of the 12 or eight sites, respectively, to contribute additively to the final coefficient. On the scale of individual plots, differences among teams or methods within each of the 120 or 270 plots, respectively, contributed additively to the final coefficient.

Results

Sampling team variability (data set 1)

The number of species recorded by the two teams was highly correlated (r = 0.95) and 44 out of 120 (36.7%) of the double sampled plots were on the line of no difference (Fig. 1). A paired *t*-test showed no difference between the two teams in species counts in the 120 plots (mean ± s.e., Team A 9.8 ± 0.5, Team B 10.0 ± 0.5; $t = -1.1$; $df = 119$; $P = 0.27$). Team records for number of taxa in a plot differed by 1.2, on average. Thirty-nine plots differed in the presence of one taxon, 20 plots differed by two taxa, 10 plots by three taxa, four plots by four taxa, two plots by five taxa and one plot by six taxa.

Sums of cover classes per plot were highly correlated for the two teams (r = 0.92), but only 16 plots (13.3%) were on the line of no difference (Fig. 1). The sum of cover classes per plot was slightly, but significantly, greater for Team B (31.7 ± 1.2) than Team A (29 ± 1.1; $t = -5.37$; $df = 119$; $P < 0.0001$). For sum of cover classes, 22 plots differed by one cover class unit, 9 plots by two units, 11 by three and four, 7 by five and six, 16 by 7, 6 by 8, 3 by 9, 4 by 10, 3 by 11 and one by 12, 13, 14, 15 and 27. There were no large differences among teams in their estimation of the cover of major growth forms (grasses/graminoids, forbs and mosses) that would indicate cover was more difficult to judge in one or more of these three general groups of plants, although Team B scored cover for forbs slightly higher than Team A (Table 1). There was a significant negative relationship between the number of occurrences of a taxon and the percentage of times cover classes differed among teams for that taxon ($r = -0.47$; $F = 37.4$; df = 132; $P < 0.0001$).

Pseudoturnover is defined as

$$(A + B)/(S_A + S_B) \times 100, \qquad (1)$$

where A and B are the number of exclusive species found by Team A and Team B, respectively, and S_A and S_B are the total number of species found by Team A and Team B, respectively (Nilsson & Nilsson 1985). The

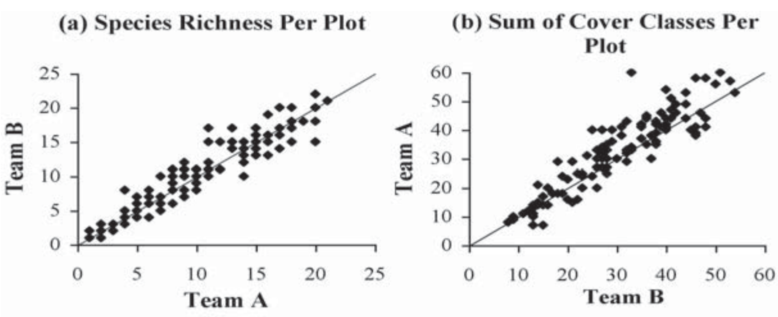

Fig. 1. Comparison of two sampling teams for (**a**) species richness per 1 m² plot and (**b**) sum of cover classes per plot. N = 120 double sampled plots. The line shown is the line of no difference.

Table 1. Summary of the number of occurrences (*N*), mean cover class (± s.e.) for forbs, grasses/graminoids and mosses as reported separately by two sampling teams in n = 120 double sampled plots.

Team	Forbs	Grasses/graminoids	Mosses
Team A	N = 791 Mean cover class 2.50 (± 0.06)	N = 367 Mean cover class 3.99 (± 0.14)	N =22 Mean cover class 1.55 (± 0.17)
Team B	N = 792 Mean cover class 2.82 (± 0.06)	N = 379 Mean cover class 3.96 (± 0.12)	N = 30 Mean cover class 2.13 (± 0.22)

overall mean pseudoturnover for double sampled plots was 19.1 ± 1.2% (n = 120). Separated by site, pseudoturnover ranged from 9.3 ± 6.5% (n = 10) at Nine Springs Meadow to a high of 30.5 ± 2.5% (n = 10) at Cherokee Ditch (Table 2). A large portion (57.7%) of pseudoturnover was attributable to taxa given the lowest cover class (< 0.5% cover). Only 3.4% of pseudoturnover was caused by taxa with cover classes six to nine (i.e. > 8% cover).

The total number of exclusive species found by each team (A + B in the pseudoturnover calculation above) was significantly, positively related to plot species richness (r = 0.81; F = 229.3; df = 119; P < 0.0001). Total plot richness also showed significant positive relationships with (a) total number of cases per plot in which the teams did not match exactly on their reporting of cover classes (r = 0.83; F = 261.8; df = 119; P << 0.0001) and (b) the total sum of cover class differences per plot (r = 0.65; F = 86.0; df = 119; P < 0.0001).

A minority of the double sampled plots (43 out of 120, 36%) was paired correctly in the 12 clustering trees (Table 3). No sites had more than six out of ten double sampled plots paired together when the nine class $\log_2$ system was used. When we pooled cover classes to obtain six or three classes, there was no significant change in the number of correctly paired double sampled plots. However, when presence/absence data were used, the number of correctly paired plots increased significantly to 90 out of 120, or 75% (Table 3; F = 9.1; df = 47; P < 0.0001).

Table 2. Mean pseudoturnover and standard error (s.e.) for the 12 sites sampled (listed alphabetically), n = 10 double sampled plots per site and the mean number of species sampled per 1 m^2 plot, based on the 30 unique plots sampled per site. Mean pseudoturnover and mean species richness m^{-2} were highly correlated (r = 0.81).

Site	% Pseudoturnover	s.e.	Mean no. spec.m^{-2}
Arboretum Pond	27.5	1.9	13.4
Cherokee Ditch	30.5	2.5	15.5
Cherokee Fen	24.5	2.5	17.7
Cherokee Marsh	15.1	3.1	7.5
Nine Springs Ditch	13.7	6.0	2.3
Nine Springs Fen	19.1	2.0	12.7
Nine Springs Meadow	9.3	6.5	2.2
Pheasant Branch Fen	28.1	4.6	14.3
Pheasant Branch Pond	20.4	2.8	10.3
Southeast Marsh	17.8	2.9	7.0
Syene Road Meadow	12.5	3.8	12.1
Wingra Marsh	11.1	2.9	5.8

Table 3. Number of double-sampled plots that were correctly paired (out of ten possible per site) in a divisive hierarchical cluster analysis using different cover class systems: (1) nine cover classes ($\log_2$; Gauch 1982); (2) six classes (see Methods); (3) three classes (see Methods) and (4) presence/absence information only. The results of a one-way ANOVA were highly significant (P < 0.0001), with more plots paired using presence/absence data than cover estimation data.

Site	9 Classes	6 Classes	3 Classes	Pres/Abs
Arboretum Pond	4	7	7	10
Cherokee Ditch	5	4	5	5
Cherokee Fen	1	2	2	5
Cherokee Marsh	0	0	1	9
Nine Springs Ditch	2	2	2	5
Nine Springs Fen	4	4	3	9
Nine Springs Meadow	6	6	7	8
Pheasant Branch Fen	5	5	10	9
Pheasant Branch Pond	2	2	3	6
Southeast Marsh	3	3	1	9
Syene Road Meadow	6	3	5	7
Wingra Marsh	5	5	3	8
Total plots paired	43	43	49	90

When the ten double sampled plots sampled on a site by a single team were combined to create a single composite vegetation for that site/team and the sites were subjected to cluster analysis, the 12 sites as sampled by Team A were correctly paired with their respective sites as sampled by Team B in the resulting tree, both when cover data and when presence/absence data were used.

Across the three analytical scales (overall, site and plot), the similarity of sampling results reported by two sampling teams using presence/absence data (Jaccard's coefficient) was 0.797, 0.693 and 0.642, respectively. The similarity of quantitative vegetation sampling results (Horn's index) for the teams was 0.894, 0.832 and 0.855 at the overall, site and plot scales, respectively.

Comparison of two sampling methods (data set 2)

The correlation between number of species per plot using the cover class method and the line intercept method was high for one line (r = 0.84) and both lines combined (r = 0.92). However, when the results for the line intercepts were plotted against the results for the cover class method, species richness values fell below the line of no difference while abundance values fell above the line of no difference (Fig. 2). The line intercept method detected significantly fewer species per plot than the cover class method, both when one line was considered (7.3 ± 0.3 for the cover class method vs 4.2 ± 0.1 for one line intercept; $t = 16.4$; $df = 267$; $P < 0.0001$) as well as when the two lines were considered together (5.4 ± 0.2 for two line intercepts; $t = 14.4$; $df = 267$; $P < 0.0001$; Fig. 2).

Correlation between the two methods for species cover estimates was high ($r = 0.79$ when one line intercept was considered and $r = 0.80$ for two lines), but again paired *t*-tests showed that species cover differed significantly for the two methods ($t = -35.5$; $df = 2347$; $P << 0.0001$), with lower mean percent cover estimates resulting for the cover class method (17.9 ± 0.6 for the cover class method vs 36 ± 0.8 for one line and 36.4 ± 0.8 for two lines; Fig. 2). When species were separated by growth form (grasses/graminoids, forbs, aquatic plants and woody plants), *t*-tests showed significant differences ($P < 0.005$) for all except woody vegetation at the overall and site scales ($P = 0.15$ and 0.058, respectively) and aquatic vegetation at the overall and site scales ($P = 0.33$ and 0.31, respectively).

Thirty out of the 270 plots sampled using cover classes were paired correctly in the clustering trees with those sampled using line intercepts. This number decreased to 25 when both lines were considered and increased to 36 when the line intercept segments were doubled in length to 20 cm each, but neither change was statistically significant ($P = 0.095$ and 0.17, respectively). However, when presence/absence data were used, the number of correct pairs increased significantly to 136, or 50% ($P = 0.005$). When the cluster analysis was carried out to evaluate sites, four of the eight sites paired correctly when cover data were used, but all eight sites paired correctly using presence/absence data.

Across the three analytical scales (overall, site and plot), the similarity of sampling results for the two sampling methods using presence/absence data (Jaccard's coefficient) was 0.752, 0.656 and 0.630, respectively. Similarity of quantitative vegetation sampling results (Horn's index) for the two methods was 0.931, 0.921 and 0.868 at the overall, site and plot scales, respectively.

Discussion

In vegetation surveys, we need to know the degree to which we can rely on differences among plots and sites as being real or due to sampling error. Although reports of quality control are quite rare in the literature of vegetation surveys we, and others, find differences attributable to sampling team and sampling method that underscore the need to adopt quality control methods.

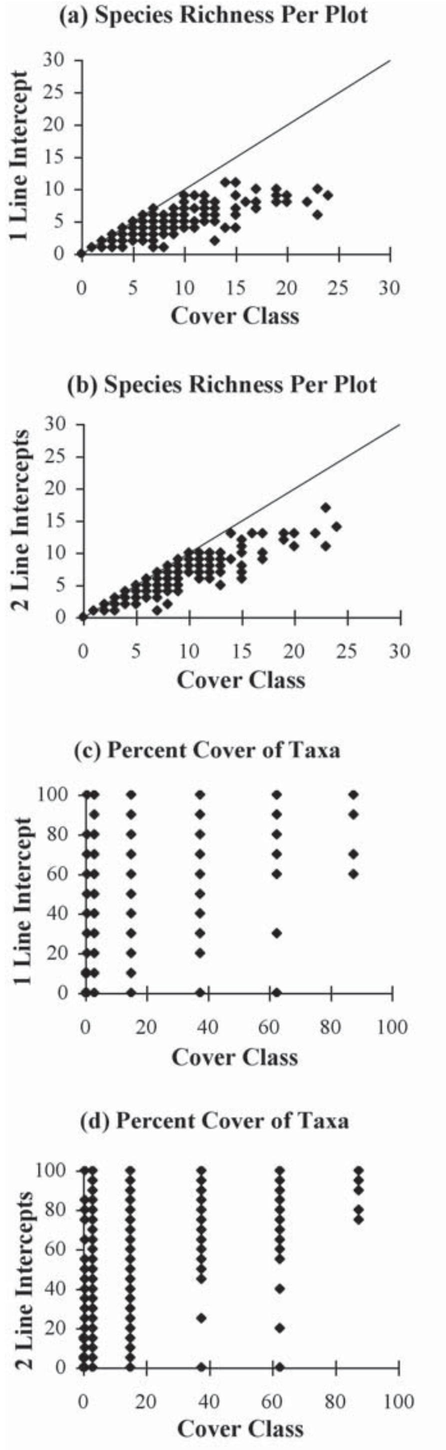

Fig. 2. Comparison of cover class and line intercept sampling methods for species richness (**a**, **b**), and percent cover of taxa present (**c**, **d**) in $N = 270$ plots. Results for the line intercept method are presented using data from one line only (**a**, **c**) or two lines (**b**, **d**) per plot. Cover class and line intercept data were transformed into percentages by taking the midpoint of each cover class and the percent of line intervals crossed by a species, respectively. The line shown is the line of no difference.

Quality control allows us to improve vegetation sampling designs by ranking/quantifying sources of variation and identifying errors and biases attributable to the specific sampling approaches used.

Sampling teams increased variability

Two well-trained, skilled sampling teams produced highly correlated results for (1) number of species per plot and (2) sum of all cover classes recorded per plot, two general descriptors of difference that are readily obtained. Few of the records were identical for species richness, but there was no apparent bias between teams for recording numbers of species. Thus, teams agreed on composite measures at the whole survey scale. At the scale of sites and individual plots, however, most plots (100 out of 120 cases) showed some pseudoturnover, and the mean overall pseudoturnover was 19.1% per plot, nearly twice that reported by Nilsson & Nilsson (1985) for two teams sampling vascular plants on 41 islands in southern Sweden (11.4%). It was also much higher than the 13% found by Lepš & Hadincová (1992) comparing two observers who sampled 25-m^2 plots in oligotrophic meadows, but similar to the 24% reported by Scott & Hallam (2003) within 0.16-m^2 plots across a range of habitat types. Since a majority of errors in our study involved taxa with the lowest cover classes, our high pseudoturnover rate could be attributable to a prevalence of rare, sterile or immature plants that were easily overlooked or misidentified (Klimeš et al. 2001). In addition, plot placement could have contributed to the high pseudoturnover since plot frames were not kept in place, but the corners of the plot were marked with flags by the first team for the second team to place their frame and sample the plot independently. Thus, slightly different boundaries could have led to the inclusion or exclusion of taxa along the 4 m of edge. If the first team leaves its sampling frame in place until the second team has sampled the plot the accuracy of pseudoturnover estimates would probably increase, but teams sampling sites independently would be more likely to produce greater differences in species lists.

A second source of variability in our study was an apparent team bias in the reporting of cover classes in the $\log_2$ system: Team B tended to score cover higher than Team A. In contrast, Lepš & Hadincová (1992) found no evidence for observer bias in the reporting of cover classes using the Braun-Blanquet method in meadows with herbaceous vegetation types similar to those in our study. There was no obvious tendency for some taxa to show greater variability in cover estimates than others, and there were no large differences among teams in their estimation of the cover within the broad categories forbs, grasses/graminoids and mosses that would indicate any of these growth forms was more difficult to estimate than the others. In contrast, Sykes et al. (1983) found that observer error was highest for taxa with thin leaves and lowest for taxa with broad leaves, although this error was not consistent within or among observers.

Many of the disagreements were for lower cover classes, which may reflect the preponderance of low cover classes in the $\log_2$ system, but disagreements also occurred for moderately abundant species, with deviations of two or more classes. Overall, there is reason to doubt that teams can agree exactly on cover class estimation. In this study, cover class estimates were made under similar conditions – teams were in the field together, sampling the same transect at the same time, under the same weather conditions and with similar degrees of fatigue. They did not consult on individual plots, but they were able to cross check intermittently. Differences would probably have been greater if these commonalities were lacking.

Variability among teams in cover estimates, as well as pseudoturnover, increased with increasing numbers of species in a plot; this demonstrates an additive effect of observer variability that can probably be ameliorated only through increasing awareness of the phenomenon by sampling crews and perhaps allocating more time and energy to species-rich plots. There was no detectable improvement in team precision over time. We attempted to relate both pseudoturnover and the number of cases in which cover classes disagreed between the two teams with the order in which sites were sampled, but we found a strong pattern in the residual plot, probably indicating site species richness was confounding the result.

The number of correctly paired double sampled plots increased from 43 to 90 out of 120, or from 36% to 75%, when we replaced cover classes with presence/absence data. This result lends some support to Wilson's (unpubl.) assertion that presence/absence data are best, but once again highlights the dual problems of (1) cover class discrepancies *and* (2) pseudoturnover among the sampling teams, since substitution of cover class data by presence/absence categories did not resolve all discrepancies. Reporting of cover classes can probably be improved only marginally. The subjectivity of cover classes was suggested as early as the 1940s (reviewed in Lepš & Hadincová 1992). Reduction of cover class categories from nine to six to three revealed no improvement in the number of correctly-paired plots, but this reduction was done *a posteriori* and results may have differed if teams had tested different cover class systems in the field. Team biases in cover estimation might be resolved with frequent, formal visual recalibration episodes before sampling each site, although we did not test this approach.

A majority of double sampled plots were not paired in the cluster analyses when cover data were used, but these many plot scale discrepancies had little bearing at the scale of whole sites. The whole site cluster analysis showed that all 12 sites paired correctly, even though the sites often represented very similar habitat types with many of the same dominant species. Thus the information on species abundance that would be lost in abandoning cover estimation methods may not be justified, particularly for vegetation studies that focus on differences among sites rather than individual plots and that employ multivariate analysis techniques. Lepš & Hadincová (1992) also found robust results using both numerical ordination and classification analyses, suggesting an insensitivity of multivariate methods to sampling errors (but see Gotfryd & Hansell 1985).

Pseudoturnover can be reduced in some cases. Klimeš et al. (2001) report substantial observer variation in the identification of small, rare, sterile plants but warn that efforts to train individuals to identify plants at all ontogenetic stages are counter productive and will not improve the error rate. Instead, they recommend that at least three observers work together and discuss all problems in the field. We found that pseudoturnover due to sampling team occasionally involved dominant or subdominant taxa as well as rare taxa. Some of the errors involving higher cover class categories were identifiable (probable field recording errors and/or misidentifications) and some were not, which makes it difficult for us to recommend a single course of action. However, in a study that spanned numerous habitat types, Scott & Hallam (2003) estimated that 5.9% of specimens were misidentified at species level and 1.9% at genus level.

Our findings highlight the difficulties of identifying some very common taxa, especially sterile specimens e.g. *Aster* ssp., *Typha* spp. and *Carex* spp. Clearly, a consensus should be reached among sampling teams on the level of identification required of difficult taxa before sampling begins, based on both the available literature and the advice of expert taxonomists. This would probably entail a reconnaissance survey of the sites to generate initial species lists before intensive quantitative sampling begins. Alternatively, agreements can be reached after sampling but before data are analysed.

Sampling methods produced different results

Two different sampling methods produced highly correlated results for species richness and cover, although there were two biases: cover classes resulted in higher species counts than line intercepts, and line intercepts produced higher estimates of cover. The use of two line intercepts instead of one barely (in the case of cover estimates) or modestly (in the case of species richness) improved the correlation between the two sampling methods. The similarity between the two methods also increased with the scale of analysis.

The high correlations between the sampling methods for general measures of composition and abundance suggest that the two sampling methods capture the same overall information about the vegetation. However, multivariate cluster analyses revealed that the cover class and line intercept methods for assessing species composition and abundance in wetlands produced different results at the plot scale most of the time. Doubling the number of transect lines did not increase the similarity with the cover class method and because abundance as measured by the line intercept was divided into more classes (ten) than the cover class method (six), we repeated the cluster analysis after combining line intercept classes to halve the number, but again this did not improve similarity of the two methods. Only reduction of cover data to presence/absence significantly increased the number of both plots and sites that were paired correctly in the cluster analyses, results which support skepticism of the use of cover estimation techniques in vegetation analyses.

When species were divided by growth form, only submerged aquatic plants and woody plants showed no significant difference in cover between the two methods. In the case of submerged aquatics, this may be because these plants occurred in few plots and when they did occur they covered large contiguous areas. Similarity between the two sampling methods is improved in this case because such widespread species are rarely missed by the line intercept method. Woody plants also occurred in few plots, but at low abundances. This would lead to greater similarity between methods because woody species missed by the line intercept method (scored as zero) would still be similar to a cover class of one.

Conclusions

We conclude that sampling team and sampling method contributed substantially to the overall variability of vegetation sampling. We recommend that all vegetation surveys incorporate quality control studies and that all members of a vegetation survey crew be trained together in field identification and sampling techniques. Multi-year studies will benefit from annual workshops to review taxa and recalibrate sampling techniques. We suggest initial site visits to generate species lists and consultation with professional taxonomists about taxa that are difficult to identify in order to reduce misidentification. When multiple observers conduct the sampling, we advise double sampling of plots and calculation of pseudoturnover as part of the quality control analysis. Plot frames should be left in place for resampling. We suggest the use of cluster analysis for comparing both sampling teams and sampling methods at the scale of plots and sites. Finally, we strongly recommend that researchers include quality control results in their publications, so that others can benefit from the knowledge gained.

Acknowledgements. This research was supported by a grant from the United States Environmental Protection Agency's Science to Achieve Results (EPA-STAR) grant number R-82801001-0 (Richard Lathrop and Ken Potter, principal investigators) and by the EPA-STAR Estuarine and Great Lakes (EaGLe) program through funding to the Great Lakes Environmental Indicators (GLEI) Project, US EPA Agreement EPA/R-8286750 (Gerald Niemi and Carol Johnston, principal investigators). Although the research described in this article has been funded wholly or in part by the US EPA, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

We thank Quentin Carpenter for his assistance developing the sampling design for the southern Wisconsin wet meadows project, and Debbie Maurer, Becky Miller, and Katy Werner for their assistance with the field sampling. Andrew Hipp taught the wetland plant identification course. We also thank Carol Johnston, Barbara Bedford, Terry Brown, and their field crews for discussion and development of methods, and Aaron Boers for field assistance in the Great Lakes wetlands project.

References

Anon. 1996. Some Basic QA/QC Concepts. In: *The volunteer monitor's guide to quality assurance project plans*, pp. 15-22. EPA 841-B-96-003, Office of Wetlands, Oceans, and Watersheds, USEPA (United States Environmental Protection Agency). Washington, DC, US.

Anon. 1998. *EPA Guidance for Quality Assurance Project Plans*. EPA 600-R-98-018, Office of Research and Development, USEPA (United States Environmental Protection Agency). Washington, DC, US.

Anon. 2000. *S-plus 2000*. MathSoft, Inc. Seattle, WA, US.

Brower, J.E., Zar, J.H. & von Ende, C.N. 1990. *Field and laboratory methods for general ecology*. 3rd. ed. Brown, Dubuque, IA, US.

Gauch, H.G. 1982. *Multivariate analyses in community ecology*. Cambridge University Press, Cambridge, UK.

Gotfryd, A. & Hansell, R.I.C. 1985. The impact of observer bias on multivariate analyses of vegetation structure. *Oikos* 45: 223-234.

Guo, Q.F. & Rundel, P.W. 1997. Measuring dominance and diversity in ecological communities: choosing the right variables. *J. Veg. Sci*. 8: 405-408.

Hope-Simpson, J.F. 1940. On the errors in the ordinary use of subjective frequency estimations in grassland. *J. Ecol*. 28: 193-209.

Kirby, K.J., Bines, T., Burn, A., MacKintosh, J., Pitkin, P. & Smith, I. 1986. Seasonal and observer differences in vascular plant records from British woodlands. *J. Ecol*. 74: 123-131.

Klimeš, L., Dančák, M., Hájek, M., Jongepierová, I. & Kučera, T. 2001. Scale-dependent biases of species counts in a grassland. *J. Veg. Sci*. 12: 699-704.

Lamacraft, R.A. 1978. Observer errors in sampling ecological variables for range condition assessments. In: Hyder, D.N. (ed.) *Proceedings of the 1st International Rangelands Congress*, pp. 514-516. Society for Rangeland Management, Denver, CO, US.

Lepš, J. & Hadincová, V. 1992. How reliable are our vegetation analyses? *J. Veg. Sci*. 3: 119-124.

Nilsson, I.N. & Nilsson, S.G. 1985. Experimental estimates of census efficiency and pseudoturnover on islands: error trend and between-observer variation when recording vascular plants. *J. Ecol*. 73: 65-70.

Oredsson, A. 2000. Choice of surveyor is vital to the reliability of assessing floristic change studies. *Watsonia* 23: 287-291.

Rich, T.C.G. & Smith, P.A. 1996. Botanical recording, distribution maps and species frequency. *Watsonia* 21: 155-167.

Rich, T.C.G. & Woodruff, E.R. 1992. Recording bias in botanical surveys. *Watsonia* 19: 73-95.

Scott, W.A. & Hallam C.J. 2003. Assessing species misidentification rates through quality assurance of vegetation monitoring. *Plant Ecol*. 165: 101-115.

Smith, A.D. 1944. A study of the reliability of range vegetation estimates. *Ecology* 25: 441-448.

Sykes, J.M., Horrill, A.D. & Mountford, M.D. 1983. Use of visual cover estimates as quantitative estimators of some British woodland taxa. *J. Ecol*. 71: 437-450.

West, N.E. & Hatton, T.J. 1990. Relative influence of observer error and plot randomization on detection of vegetation change. *Coenoses* 5: 45-49.

Received 31 January 2003;
Revision received 2003;
Accepted 11 June 2003.
Final version received 5 August 2003;
Co-ordinating Editor: K. Thompson.